# Continuum Limits of Markov Chains with Application to Network Modeling

Yang Zhang, *Student Member, IEEE,* Edwin K. P. Chong, *Fellow, IEEE,* Jan Hannig, and Donald Estep



***Abstract*—In this paper we investigate the continuum limits of a class of Markov chains. The investigation of such limits is motivated by the desire to model very large networks. We show that under some conditions, a sequence of Markov chains converges in some sense to the solution of a partial differential equation. Based on such convergence we approximate Markov chains modeling networks with a large number of components by partial differential equations. While traditional Monte Carlo simulation for very large networks is practically infeasible, partial differential equations can be solved with reasonable computational overhead using well-established mathematical tools.**

***Index Terms*—Continuum modeling, Markov chain, partial differential equation, large network modeling, wireless sensor network.**

## I. Introduction

NETWORK modeling is an important tool in the analysis and design of networks. Many network characteristics of interest can be modeled by Markov chains, where Monte Carlo simulation has been the traditional approach [1]. With the enormous growth in the size and complexity of today's networks, their simulation becomes more computationally expensive in both time and hardware. Some effort has been made to exploit the computing powers of distributed computer networks, such as parallel simulation techniques, where the number of processors needed in the simulation increases with the number of nodes in the network [2], [3]. However, for networks involving a very large number of nodes, Monte Carlo simulation eventually becomes practically infeasible.

In this paper we address this problem by focusing on the global characteristics of an entire network rather than those of its individual components. The idea is to approximate the underlying Markov chain modeling a certain network characteristic by a partial differential equation (PDE).

As a concrete familiar example, which we present in Section II, consider multiple i.i.d. (independent and identically distributed) random walks of $M$ particles on a network consisting of $N$ points. For any vector $x$, let $x^T$ be its transpose. Let the Markov chain modeling the network characteristic be $X_N(k) = [X_N(k,1),\dots,X_N(k,N)]^T \in \mathbb{R}^N$, where $X_N(k,n)$ is the number of particles at point $n$ at time $k$. If we treat $N$ and $M$ as indices that grow, this defines a family of Markov chains indexed by $N$ and $M$. We show that as $M \to \infty$ and $N \to \infty$, $X_N(k)$ converges in some sense to its continuum limit, a deterministic function with continuous time and space variables. Under certain conditions, it is possible to characterize such a function as the solution of a PDE [4]–[6]. This itself is not a new result, but helps to illustrate our aim.

Indeed, our development here is motivated by the network modeling strategy in [7] and the need for a rigorous description of its underlying limiting process. We illustrate in Section III the convergence of the sequence of Markov chains to the PDE in a two-step procedure. Suppose the evolution of $X_N(k)$ is governed by a certain stochastic difference equation with a "normalizing" parameter $M$. Let $x_N(k)$ be the normalized deterministic sequence governed by the corresponding "expected" and deterministic difference equation. First, we show in Section III-B that $X_N(k)/M$ is close to $x_N(k)$, in the sense that as $M \to \infty$, both their continuous-time extensions converge to the solution of an ordinary differential equation (ODE). Second, we show in Section III-C that as $N \to \infty$, $x_N(k)$ converges to the solution of a PDE. Therefore, as $M \to \infty$ and $N \to \infty$, $X_N(k)/M$ converges to the PDE solution.

Our procedure provides an approach to approximating Markov chains that model large networks by PDEs. PDEs are widely used to formulate time-space phenomena in physics, chemistry, ecology, and economics (e.g., [8]–[11]), and there are well-established mathematical tools for solving them such as Matlab and Comsol, which use finite element method [12] or finite difference method [13]. In contrast to Monte Carlo simulation, our approach enables us to use these tools to greatly reduce computation time, which makes it possible to carry out the analysis, design, and optimization for very large networks. We present in Section IV an example of the application of our approach to the modeling of a large wireless sensor network. In this example, we derive an explicit nonlinear diffusion-convection PDE, whose solution captures the dynamic behavior of the data message queues in the network. We show that although the PDE approximation takes only a tiny fraction of the computation time of the Monte Carlo simulation, there is a strong agreement between their simulation results.

Continuum modeling has been well-established in fields such as physics, mechanics, transportation, and biology (e.g., [14]–[17]). Its applications in communication networks, how-

This research was supported in part by NSF grant ECCS-0700559.

Yang Zhang and Edwin K. P. Chong are with the Department of Electrical and Computer Engineering, Colorado State University, Ft. Collins, CO 80523-1373 yzhangcn@mail.engr.colostate.edu & edwin.chong@colostate.edu

Jan Hannig is with the Department of Statistics and Operation Research, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3260 jan.hannig@unc.edu

Donald Estep is with the Department of Mathematics and Department of Statistics, Colorado State University, Fort Collins, CO 80523-1373 estep@math.colostate.edu

A preliminary version of parts of the work of this paper was presented at the 49th IEEE Conference on Decision and Control.

ever, are relatively new and rare. Among these, to our best knowledge, our approach is the first to address the time-space characteristics of communication networks with a large number of nodes. In contrast, for example, [18]–[20] deal with networks with heavy traffic instead of large number of nodes; [21], [22] present scaling laws of the network traffic without characterizing the actual traffic over time and space; and [23], [24], which use mean field methods, only keep track of the statistical features of the networks such as the fraction of nodes in each network state.

## II. Continuum Limit of Multiple Random Walks

In this section we present an illustrative example of approximating multiple i.i.d. random walks by a PDE. First consider a single random walk on a one-dimensional network consisting of $N$ points uniformly placed over $\mathcal{D} = [0,1]$, as shown in Fig. 1. Hence the distance between two neighboring point is $ds = 1/(N+1)$. At each time instant, the particle at point $n$, where $n = 1,\dots,N$, randomly chooses to move to its left or right neighboring point with probability $P_r(n)$ and $P_l(n)$, respectively. Let the length between two time instants be $dt = 1/M$. We set $dt = ds^2$, which is a standard time-space scaling approach to ensuring the convergence of the difference equation to a PDE. We assume a "sink" boundary condition, i.e., the particle vanishes when it reaches the ends of $\mathcal{D}$ (though "walls" at the boundary are equally treatable).

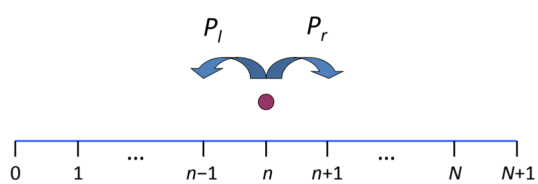


Fig. 1. An illustration of a one-dimensional single random walk.

Now consider $M$ random walks on the same network, where the particle in each random walk behaves independently identically as in the single random walk described above. Let $B_i(k,n)$ be the Bernoulli random variable representing the presence of the $i$th particle at point $n$ at time instant $k$, where $k = 0,1,\dots$ and $n = 1,\dots,N$. Define $B_i(k) = [B_i(k,1),\dots,B_i(k,N)]^T \in \mathbb{R}^N$. According to the behavior of the particle in the single random walk, for $i = 1,\dots,M$,

$$\begin{aligned} &B_i(k+1,n) - B_i(k,n) \\ &= \begin{cases} B_i(k,n-1), & \text{with probability } P_r(n-1); \\ B_i(k,n+1), & \text{with probability } P_l(n+1); \\ -B_i(k,n), & \text{with probability } P_r(n)+P_l(n); \\ 0, & \text{otherwise,} \end{cases} \end{aligned}$$

where $B_i(k,n)$ with $n \le 0$ or $n \ge N+1$ are defined to be zero. Let the function $F_N(x,U(k))$, where $U(k)$ are i.i.d. and do not depend on $x$, be such that for $i = 1,\dots,M$,

$$B_i(k+1) = B_i(k) + F_N(B_i(k),U(k)). \tag{1}$$

Then for $x = [x_1,\dots,x_N]^T$, the $n$th component of $F_N(x,U(k))$, where $n = 1,\dots,N$, is

$$\begin{cases} x_{n-1}, & \text{with probability } P_r(n-1); \\ x_{n+1}, & \text{with probability } P_l(n+1); \\ -x_n, & \text{with probability } P_r(n)+P_l(n); \\ 0, & \text{otherwise,} \end{cases} \tag{2}$$

where $x_n$ with $n \le 0$ or $n \ge N+1$ are defined to be zero.

Let $X_N(k,n)$ be the number of particles at point $n$ at time $k$. Then

$$X_N(k,n) = \sum_{i=1}^{M} B_i(k,n). \tag{3}$$

Define $X_N(k) = [X_N(k,1),\dots,X_N(k,N)]^T$, which forms a discrete-time Markov chain with state space $\mathbb{R}^N$. Since $F_N$ is linear, it follows from (3) that

$$X_N(k+1) = X_N(k) + F_N(X_N(k),U(k)).$$

Let

$$f_N(x) = EF_N(x,U(k)), \quad x \in \mathbb{R}^N.$$

It follows from (2) that for $x = [x_1,\dots,x_N]^T$, the $n$th component of $f_N(x)$, where $n = 1,\dots,N$, is

$$P_r(n-1)x_{n-1} + P_l(n+1)x_{n+1} - (P_r(n)+P_l(n))x_n, \tag{4}$$

where $x_n$ with $n \le 0$ or $n \ge N+1$ are defined to be zero. By (1) and the linearity of $F_N$, for $i = 1,\dots,M$,

$$EB_i(k+1) = EB_i(k) + f_N(EB_i(k)). \tag{5}$$

Notice that, since the random walks are i.i.d., $EB_i$ does not depend on $i$. Define a deterministic sequence $x_N(k)$ by

$$x_N(k+1) = x_N(k) + f_N(x_N(k)), \tag{6}$$

where

$$x_N(0) = \frac{X_N(0)}{M}, \text{ a.s. (almost surely).} \tag{7}$$

We seek to approximate $X_N(k)$ by a continuum model, where the time and space indices $k$ and $n$ are made continuous as $N \to \infty$ and $M \to \infty$ in the following two steps: First, define

$$X_{oN}(\tilde{t}) = \frac{X_N(\lfloor M\tilde{t} \rfloor)}{M},$$

the continuous-time extension of $X_N(k)$ by piecewise-constant time extensions with interval length $dt = 1/M$ and scaled by $1/M$. Second, define $X_{pN}(t,s)$ to be the continuous-space extension of $X_{oN}(\tilde{t})$ by piecewise-constant space extensions on $\mathcal{D}$ with interval length $ds$. Notice that as $N \to \infty$, $ds \to 0$. Thus $X_{pN}$ is the continuous-time-space extension of $X_N(k)$. Similarly, define $x_{oN}(\tilde{t}) = x_N(\lfloor M\tilde{t} \rfloor)$, the piecewise-constant continuous-time extension of $x_N(k)$, and $x_{pN}(t,s)$, the piecewise-constant continuous-space extension of $x_{oN}(\tilde{t})$. Thus $x_{pN}$ is the continuous-time-space extension of $x_N(k)$.

Now we show that for $M$ sufficiently large, $X_{pN}$, the continuous-time-space extension of $X_N(k)$, is close to $x_{pN}$, the continuous-time-space extension of $x_N(k)$. By (3) and the strong law of large numbers (SLLN), for each $k$,

$$\lim_{M \to \infty} \frac{X_N(k)}{M} = EB_i(k) \text{ a.s.}$$

By this and (7),

$$\lim_{M\to\infty} x_N(0) = EB_i(0) \text{ a.s.}$$

By (5) and (6), $x_N(k)$ and $EB_i(k)$ satisfy the same difference equation. Then we have for each $k$,

$$\lim_{M\to\infty} x_N(k) = EB_i(k) \text{ a.s.}$$

Hence for each $k$,

$$\lim_{M\to\infty} \frac{X_N(k)}{M} = x_N(k) \text{ a.s.}$$

Therefore, $X_{oN}$ and $x_{oN}$ are close for large $M$ in the sense that

$$\lim_{M\to\infty} \|X_{oN}(\tilde{t}) - x_{oN}(\tilde{t})\|_\infty^{(N)} = 0 \text{ a.s.,} \tag{8}$$

where $\|\cdot\|_\infty^{(N)}$ is the $\infty$-norm on $\mathbb{R}^N$. Note that

$$\|X_{pN}(\cdot,t) - x_{pN}(\cdot,t)\|_\infty^{(\mathcal{D})} = \|X_{oN} - x_{oN}\|_\infty^{(N)},$$

where $\|\cdot\|_\infty^{(\mathcal{D})}$ is the $\infty$-norm on $\mathbb{R}^{\mathcal{D}}$, the space of functions of $\mathcal{D}\to\mathbb{R}$. Then by (8), $X_{pN}$ and $x_{pN}$ are close to each other for large $M$ in the sense that

$$\lim_{M\to\infty} \|X_{pN}(\cdot,t) - x_{pN}(\cdot,t)\|_\infty^{(\mathcal{D})} = 0 \text{ a.s.} \tag{9}$$

Therefore, we can approximate $X_{pN}$ by $x_{pN}$ for $M$ sufficiently large.

Next we show that as $N\to\infty$, $x_{pN}$ satisfies a certain PDE that is easily solvable. By (4) we have for $n = 1,\dots,N$,

$$\begin{aligned} &x_N(k+1,n) - x_N(k,n)\\ &= P_r(n-1)x_N(k,n-1) + P_l(n+1)x_N(k,n+1)\\ &\quad - (P_r(n)+P_l(n))x_N(k,n), \end{aligned}$$

where $x_N(k,n)$ with $n\le 0$ or $n \ge N+1$ are defined to be zero. Assume $P_l(n) = p_l(nds)$ and $P_r(n) = p_r(nds)$, where $p_l(s)$ and $p_r(s)$ are real-valued functions defined on $\mathcal{D}$. Then by the definition of $x_{pN}$, it follows that for $s\in\mathcal{D}$ and $t>0$,

$$\begin{aligned} &x_{pN}(t+dt,s) - x_{pN}(t,s)\\ &= p_r(s-ds)x_{pN}(t,s-ds) + p_l(s+ds)x_{pN}(t,s+ds)\\ &\quad - (p_r(s)+p_l(s))x_{pN}(t,s). \end{aligned} \tag{10}$$

To ensure a finite non-degenerate limit, we assume

$$p_l(s) = b(s) + c_l(s)ds \text{ and } p_r(s) = b(s) + c_r(s)ds.$$

Define $c = c_l - c_r$. We call $b$ the diffusion coefficient and $c$ the convection coefficient, for a greater $b$ means more rapid diffusion and a greater $c$ means a larger directional bias. Assume that $b\in\mathcal{C}^2$ and $c\in\mathcal{C}^1$. Assume that $x_{pN}$ is twice continuously differentiable in $s$. Put into (10) the Taylor expansions

$$\begin{aligned} x_{pN}(t,s\pm ds) &= x_{pN}(t,s) \pm \frac{\partial x_{pN}}{\partial s}(t,s)ds\\ &\quad + \frac{\partial^2 x_{pN}}{\partial s^2}(t,s)\frac{ds^2}{2} + o(ds^2), \end{aligned} \tag{11}$$

$$b(s\pm ds) = b(s) \pm b_s(s)ds + b_{ss}(s)\frac{ds^2}{2} + o(ds^2), \tag{12}$$

and

$$c(s\pm ds) = c(s) \pm c_s(s)ds + o(ds), \tag{13}$$

where a single subscript $s$ represents first derivative and a double subscript $ss$ represents second derivative. Then we have

$$\begin{aligned} &x_{pN}(t+dt,s) - x_{pN}(t,s) = b(s)\frac{\partial^2 x_{pN}}{\partial s^2}(t,s)ds^2\\ &+ (2b_s(s)+c(s))\frac{\partial x_{pN}}{\partial s}(t,s)ds^2\\ &+ (b_{ss}(s)+c_s(s))x_{pN}(t,s)ds^2 + o(ds^2). \end{aligned} \tag{14}$$

Divide both sides of (14) by $dt = ds^2$ and get

$$\begin{aligned} &\frac{x_{pN}(t+dt,s) - x_{pN}(t,s)}{dt}\\ &= b(s)\frac{\partial^2 x_{pN}}{\partial s^2}(t,s) + (2b_s(s)+c(s))\frac{\partial x_{pN}}{\partial s}(t,s)\\ &\quad + (b_{ss}(s)+c_s(s))x_{pN}(t,s) + \frac{o(ds^2)}{ds^2}. \end{aligned}$$

As $N\to\infty$, $ds\to 0$, and hence $dt = ds^2 \to 0$. Assume that $x_{pN}$ is continuously differentiable in $t$. Then by taking the limit as $N\to\infty$ and rearranging, we get a PDE that $x_{pN}$ satisfies:

$$\begin{aligned} \dot{x}_{pN}(t,s) &= \frac{\partial}{\partial s}\left(b(s)\frac{\partial x_{pN}}{\partial s}(t,s)\right)\\ &\quad + \frac{\partial}{\partial s}((b_s(s)+c(s))x_{pN}(t,s)), \end{aligned}$$

for $t>0$ and $s\in\mathcal{D}$, with boundary condition $x_{pN}(t,s) = 0$.

As $N\to\infty$, $dt = ds^2 \to 0$, and hence $M = 1/dt = 1/ds^2 \to \infty$. Then by (9), for $N$ sufficiently large, $X_{pN}$, the continuous-time-space extension of $X_N(k)$, is close to $x_{pN}$, the continuous-time-space extension of $x_N(k)$. Therefore, we can approximate $X_N(k)$ by the solution of the above PDE called the one-dimensional diffusion-convection equation, which can be easily solved [25]. Note that our derivation here differs from that of the well-studied Fokker-Planck equation (also known as the Kolmogorov forward equation) [26], whereas the latter originates from the study of the probability density of a Wiener process.

This motivational example raises some questions that must be answered by the convergence analysis of the underlying limiting process. First, general networks may exhibit more complex behaviors. For example, $F_N$ might no longer be linear; and SLLN might not apply in many scenarios since node behaviors are not necessarily i.i.d. Specifically, the analysis above does not apply to the network Markov chain in [7]. To find the conditions under which (8) holds in more general setting, in Section III-B we apply Kushner's weak convergence theorem in [4] to a more general class of systems modeled by Markov chains. Moreover, we need to show in what sense and under what conditions $X_{pN}$ converges to the solution of the PDE. We analyze such convergence and provide its sufficient conditions in Section III-C.

## III. Continuum Limits of Markov Chains

In this section we analyze the convergence of a sequence of Markov chains to the solution of a PDE in a two-step procedure. We provide sufficient conditions for this convergence.

### A. General Setting

Consider $N$ points placed over a Euclidean domain $\mathcal{D}$ representing a spatial region. We assume that these points form a *uniform* grid, though our approach can later be generalized to nonuniform cases. We will refer to these $N$ points in $\mathcal{D}$ as *grid* points and denote the distance between any two neighboring grid points by $ds_N$.

Consider a discrete-time Markov chain

$$X_N(k) = [X_N(k,1), \dots, X_N(k,N)]^T \tag{15}$$

with state space $\mathbb{R}^N$. Here $X_N(k,n)$ is the real-valued state of point $n$ at time $k$, where $n = 1, \dots, N$ is a *spatial* index and $k = 0, 1, \dots$ is a *temporal* index.

Suppose that the evolution of $X_N(k)$ is described by the stochastic difference equation

$$X_N(k+1) = X_N(k) + F_N(X_N(k)/M, U(k)), \tag{16}$$

where $U(k)$ are i.i.d. and do not depend on the state $X_N(k)$, $M$ is a "normalizing" parameter, and $F_N$ is a given function. Let

$$f_N(x) = EF_N(x, U(k)), \quad x \in \mathbb{R}^N. \tag{17}$$

Define a deterministic sequence $x_N(k)$ by

$$x_N(k+1) = x_N(k) + \frac{1}{M} f_N(x_N(k)), \tag{18}$$

where $x_N(0) = X_N(0)/M$ a.s. In the next subsection, we show that under certain conditions, $X_N(k)/M$ and $x_N(k)$ are close in some sense.

### B. Convergence to ODE

Let $X_{oN}(\tilde{t})$ be the continuous-time extension of $X_N(k)$ by piecewise-constant time extensions with interval length $1/M$ and scaled by $1/M$, i.e., for arbitrary $\tilde{t} \in \mathbb{R}$,

$$X_{oN}(\tilde{t}) = X_N(\lfloor M\tilde{t} \rfloor)/M. \tag{19}$$

It follows that for each $k$, $X_{oN}(k/M) = X_N(k)/M$. Similarly we define $x_{oN}(\tilde{t})$, the continuous-time extension of $x_N(k)$ by

$$x_{oN}(\tilde{t}) = x_N(\lfloor M\tilde{t} \rfloor). \tag{20}$$

For fixed $\tilde{T}_N > 0$, let $D^N[0, \tilde{T}_N]$ be the space of $\mathbb{R}^N$-valued Càdlàg functions on $[0, \tilde{T}_N]$, i.e., functions that are right-continuous at each $t \in [0, \tilde{T}_N)$ and have left-hand limits at each $t \in (0, \tilde{T}_N]$. As defined in (19) and (20) respectively, both $X_{oN}(\tilde{t})$ and $x_{oN}(\tilde{t})$ with $\tilde{t} \in [0, \tilde{T}_N]$ are in $D^N[0, \tilde{T}_N]$. Since both $X_{oN}(\tilde{t})$ and $x_{oN}(\tilde{t})$ depend on $M$, each one of them forms a sequence of functions in $D^N[0, \tilde{T}_N]$ indexed by $M = 1, 2, \dots$.

Define the $\infty$-norm $\|\cdot\|_\infty^{(o)}$ on $D^N[0, \tilde{T}_N]$, i.e., for $x \in D^N[0, \tilde{T}_N]$,

$$\|x\|_\infty^{(o)} = \max_{n=1,\dots,N} \sup_{t \in [0, \tilde{T}_N]} |x^n(t)|,$$

where $x^n$ is the $n$th components of $x$. A sequence of functions $x_M \in D^N[0, \tilde{T}_N]$ is said to converge uniformly to a function $x \in D^N[0, \tilde{T}_N]$ if as $M \to \infty$, $\|x_M - x\|_\infty^{(o)} \to 0$. In this paper, we use the notation "$\Rightarrow$" for weak convergence and "$\xrightarrow{P}$" for convergence in probability.

Let $f_N$ be defined as in (17). Now we present a lemma stating that under some conditions, as $M \to \infty$, $X_{oN}$ converges uniformly to a limiting function $y$, the solution of the ODE $\dot{y} = f_N(y)$, on $[0, \tilde{T}_N]$, and $X_{oN}$ converges uniformly to the same solution on $[0, \tilde{T}_N]$.

*Lemma 1:* Assume:

(1a) There exists an identically distributed sequence $\{\lambda(k)\}$ of integrable random variables such that for each $k$ and $x$, $|F_N(x, U(k))| \le \lambda(k)$ a.s.;

(1b) the function $F_N(x, U(k))$ is continuous in $x$ a.s.; and

(1c) the ODE $\dot{y} = f_N(y)$ has a unique solution on $[0, \tilde{T}_N]$ for any initial condition $y(0)$.

Suppose that as $M \to \infty$,

$$X_{oN}(0) \xrightarrow{P} y(0) \text{ and } x_{oN}(0) \to y(0).$$

Then, as $M \to \infty$,

$$\|X_{oN} - y\|_\infty^{(o)} \xrightarrow{P} 0 \text{ and } \|x_{oN} - y\|_\infty^{(o)} \to 0$$

on $[0, \tilde{T}_N]$, where $y$ is the unique solution of $\dot{y} = f_N(y)$ with initial condition $y(0)$.

To prove Lemma 1, we first present a lemma on weak convergence due to Kushner [4].

*Lemma 2:* Assume:

(2a) The set

$$\{|F_N(x, U(k))| : k \ge 0\}$$

is uniformly integrable;

(2b) for each $k$ and each bounded random variable $X$,

$$\lim_{\delta \to 0} E \sup_{|Y| \le \delta} |F_N(X, U(k)) - F_N(X+Y, U(k))| = 0;$$

and

(2c) there is a function $\hat{f}_N(\cdot)$ [continuous by (b)] such that as $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{n} F_N(x, U(k)) \xrightarrow{P} \hat{f}_N(x).$$

Suppose that $\dot{y} = \hat{f}_N(y)$ has a unique solution on $[0, \tilde{T}_N]$ for each initial condition, and that $X_{oN}(0) \Rightarrow y(0)$. Then as $M \to \infty$,

$$\|X_{oN} - y\|_\infty^{(o)} \Rightarrow 0 \text{ on } [0, \tilde{T}_N].$$

We note that in Kushner's work, the convergence of $X_{oN}$ to $y$ is stated in terms of Skorokhod norm [4], but it is equivalent to the $\infty$-norm in our case where the functions are defined on finite time intervals [27].

We now prove Lemma 1 by showing that the assumptions (2a)–(2c) in Lemma 2 hold under the assumptions (1a)–(1c) in Lemma 1.

*Proof of Lemma 1:*

1) Since $\lambda(k)$ is integrable, as $a \to \infty$,

$$E|\lambda(k)|1_{\{|\lambda(k)|>a\}} \to 0,$$

where $1_A$ is the indicator function of set $A$. By Assumption (1a), for each $k$ and $x$,

$$F_N(x, U(k)) \le \lambda(k) \text{ a.s.}$$

Therefore for each $x$ and $a > 0$,

$$\begin{aligned} &E|F_N(x, U(k))|1_{\{|F_N(x,U(k))|>a\}} \\ &\le E|\lambda(k)|1_{\{|F_N(x,U(k))|>a\}} \\ &\le E|\lambda(k)|1_{\{|\lambda(k)|>a\}}. \end{aligned}$$

Hence as $a \to \infty$,

$$\sup_{k\ge 0} E|F_N(x, U(k))|1_{\{|F_N(x,U(k))|>a\}} \to 0,$$

i.e., the family $\{|F_N(x, U(k))| : k \ge 0\}$ is uniformly integrable and Assumption (2a) holds.

2) By Assumption (1b), $F_N(x, U(k))$ is continuous in $x$ a.s. Then for each bounded $X$ and each $k$,

$$\lim_{\delta\to 0} \sup_{|Y|\le\delta} |F_N(X, U(k)) - F_N(X + Y, U(k))| = 0 \text{ a.s.}$$

By Assumption (1a), for each $x$ and each $k$, there exists an integrable random variable $\lambda(k)$ such that $|F_N(x, U(k))| \le \lambda(k)$ a.s. It follows that for each bounded $X$, each $k$, and each $Y$ such that $|Y| \le \delta$,

$$\begin{aligned} &|F_N(X, U(k)) - F_N(X + Y, U(k))| \\ &\le |F_N(X, U(k))| + |F_N(X + Y, U(k))| \le 2\lambda(k). \end{aligned}$$

Hence for each $\delta$,

$$\left| \sup_{|Y|\le\delta} |F_N(X, U(k)) - F_N(X + Y, U(k))| \right| \le 2\lambda(k),$$

an integrable random variable. By the dominant convergence theorem,

$$\begin{aligned} &\lim_{\delta\to 0} E \sup_{|Y|\le\delta} |F_N(X, U(k)) - F_N(X + Y, U(k))| \\ &= E \lim_{\delta\to 0} \sup_{|Y|\le\delta} |F_N(X, U(k)) - F_N(X + Y, U(k))| \\ &= 0. \end{aligned}$$

Hence Assumption (2b) holds.

3) Since $U(k)$ are i.i.d., by the weak law of large numbers and the definition of $f_N$ in (17), as $n \to \infty$,

$$\frac{1}{n} \sum_{k=0}^{n} F_N(x, U(k)) \xrightarrow{P} f_N(x).$$

Hence Assumption (2c) holds.

Then, by Lemma 2, as $M \to \infty$, $\|X_{oN} - y\|_\infty^{(o)} \Rightarrow 0$ on $[0, \tilde{T}_N]$. For each sequence of random processes $\{X_n\}$, if $A$ is a constant, $X_n \Rightarrow A$ if and only if $X_n \xrightarrow{P} A$. Therefore, as $M \to \infty$, $\|X_{oN} - y\|_\infty^{(o)} \xrightarrow{P} 0$ on $[0, \tilde{T}_N]$. The same argument implies the deterministic convergence of $x_{oN}$: as $M \to \infty$, $\|x_{oN} - y\|_\infty^{(o)} \to 0$ on $[0, \tilde{T}_N]$. ∎

Based on Lemma 1, we get the following lemma, which states that $X_{oN}$ and $x_{oN}$ are close with high probability when $M$ is large.

*Lemma 3:* Let the assumptions in Lemma 1 hold. Then for any sequence $\{\zeta_N\}$, for each $N$, and for $M$ sufficiently large, we have

$$P\{\|X_{oN} - x_{oN}\|_\infty^{(o)} > \zeta_N\} \le 1/N^2 \text{ on } [0, \tilde{T}_N].$$

*Proof:* By Lemma 1, for each $N$, as $M \to \infty$,

$$\|X_{oN} - y\|_\infty^{(o)} \xrightarrow{P} 0 \text{ and } \|x_{oN} - y\|_\infty^{(o)} \to 0 \text{ on } [0, \tilde{T}_N].$$

By the triangle inequality

$$\|X_{oN} - x_{oN}\|_\infty^{(o)} \le \|X_{oN} - y\|_\infty^{(o)} + \|x_{oN} - y\|_\infty^{(o)},$$

it follows that as $M \to \infty$, $\|X_{oN} - x_{oN}\|_\infty^{(o)} \xrightarrow{P} 0$ on $[0, \tilde{T}_N]$. This finishes the proof. ∎

Since $X_{oN}$ and $x_{oN}$ are the piecewise continuous-time extensions of $X_N$ and $x_N$ by constant interpolation, respectively, we have the following corollary.

*Corollary 1:* Fix $\tilde{T}_N$ and let $\tilde{K}_N = \lfloor \tilde{T}_N M \rfloor$. Let the assumptions in Lemma 1 hold. Then for any sequence $\{\zeta_N\}$, for each $N$, and for $M$ sufficiently large, we have

$$P\left\{ \max_{\substack{k=0,\dots,\tilde{K}_N \\ n=1,\dots,N}} \left| \frac{X_N(k,n)}{M} - x_N(k,n) \right| > \zeta_N \right\} \le \frac{1}{N^2}.$$

We use Lemma 3 and Corollary 1 in the next subsection.

### C. Convergence to PDE

In the last subsection, we stated conditions under which the continuous-time extensions of $X_N(k)$ and $x_N(k)$ are close asymptotically (as $M \to \infty$) with high probability. In this subsection, we further let $N \to \infty$ and state conditions under which $x_N(k)$ is close asymptotically to the solution of a PDE. This leads to the convergence of $X_N(k)/M$ to the PDE solution as $M \to \infty$ and $N \to \infty$.

Assume that the domain $\mathcal{D}$ introduced in Section III-A is compact and convex, and let $w : \mathcal{D} \to \mathbb{R}$ be in $\mathcal{C}^2$. Given a fixed $N$, let $V_N$ be the set of the $N$ grid points in $\mathcal{D}$. Let $y_N$ be the vector in $\mathbb{R}^N$ composed of the values of $w$ at the grid points $v_N(n) \in V_N$, $n = 1, \dots, N$, i.e., $y_N = [w(v_N(1)), \dots, w(v_N(N))]^T$.

Given $s \in \mathcal{D}$, let $\{s_N\} \subset \mathcal{D}$ be a sequence of grid points in $\mathcal{D}$ such that as $N \to \infty$, $s_N \to s$, where for each $N$, $s_N$ is a grid point in $V_N$. Let $f_N(y_N, s_N)$ be the component of the vector $f_N(y_N)$ corresponding to the location $s_N$. For example, for $N = 5$, if $s_5 = v_5(4)$ in $V_5$, then $f_5(y_5, s_5)$ is the 4th component of the vector $f_5(y_5)$.

Assume that there exist sequences $\{\delta_N\}$, $\{\beta_N\}$, $\{\gamma_N\}$, and $\{\rho_N\}$, functions $f$ and $h$, and $0 < c < \infty$, such that as $N \to \infty$, $\delta_N \to 0$, $\delta_N/\beta_N \to 0$, $\gamma_N \to 0$, $\rho_N \to 0$, and:

- for any $s_N$ such that $s_N \to s$, where $s$ is in the interior of $\mathcal{D}$, there exists a sequence of functions $\phi_N : \mathcal{D} \to \mathbb{R}$ such that

$$\begin{aligned} f_N(y_N, s_N)/\delta_N = f(s_N, w(s_N), \nabla w(s_N), \nabla^2 w(s_N)) \\ + \phi_N(s_N), \end{aligned} \tag{21}$$

and for $N$ sufficiently large,

$$|\phi_N(s_N)| \le c\gamma_N; \tag{22}$$

and

- for any $s_N$ such that $s_N \to s$, where $s$ is on the boundary of $\mathcal{D}$, there exists a sequence of functions $\varphi_N : \mathcal{D} \to \mathbb{R}$ such that

$$f_N(y_N, s_N)/\beta_N = h(s_N, w(s_N), \nabla w(s_N), \nabla^2 w(s_N))$$

$$+ \varphi_N(s_N), \tag{23}$$

and for $N$ sufficiently large, $|\varphi_N(s_N)| \le c\rho_N$.

Here, $\nabla^i w$ represents all the $i$th order derivatives of $w$, where $i = 1, 2$.

These assumptions are technical conditions on the asymptotic behavior of the sequence of functions $f_N$. The basic idea is that $f_N(y_N, s_N)$ is asymptotically close to some function of terms that look like the right-hand side of a time-dependent PDE. Typically, checking these conditions amounts to simply an algebraic exercise. A concrete example of this is given in the next section.

The basic idea underlying the analysis in the remainder of this subsection is this. Recall that $x_N(k)$ is defined by (18). Suppose we associate the discrete time $k$ with points on the real line spaced apart by a distance proportional to $\delta_N$. Then, the above technical assumption implies that $x_N(k)$ is, in some sense, close to the solution of a PDE of the form $\dot{z} = f(s, z, \nabla z, \nabla^2 z)$ with boundary condition $h(s, z, \nabla z, \nabla^2 z) = 0$. Because the Markov chain $X_N(k)/M$ is close to $x_N(k)$, as established in the last subsection, it is also close to the solution of the PDE. The remainder of this subsection is devoted to developing this argument rigorously.

Fix $T > 0$. Assume that there exists a function $z : [0, T] \times \mathcal{D} \to \mathbb{R}$ that solves the PDE

$$\dot{z}(t, s) = f(s, z(t, s), \nabla z(t, s), \nabla^2 z(t, s)), \tag{24}$$

with boundary condition

$$h(s, z(t, s), \nabla z(t, s)\nabla^2 z(t, s)) = 0$$

and initial condition $z(0, s) = z_0(s)$. Here, $\nabla^i z(t, s)$ represents all the $i$th order partial derivatives of $z(t, s)$ with respect to $s$, where $i = 1, 2$.

Define

$$dt_N = \delta_N / M. \tag{25}$$

Define

$$K_N = \lfloor T/dt_N \rfloor \text{ and } t_N(k) = k dt_N.$$

Define

$$z_N(k, n) = z(t_N(k), v_N(n))$$

and let $z_N(k) = [z_N(k, 1), \dots, z_N(k, N)]^T \in \mathbb{R}^N$.

Denote the $\infty$-norm on $\mathbb{R}^N$ by $\|\cdot\|_\infty^{(N)}$. That is, for $x \in \mathbb{R}^N$, with the $n$th element being $x(n)$,

$$\|x\|_\infty^{(N)} = \max_{1 \le n \le N} |x(n)|.$$

Denote the $\infty$-norm on $\mathbb{R}^{N \times K_N}$ also by $\|\cdot\|_\infty^{(N)}$. That is, for $x = [x(1), \dots, x(K_N)] \in \mathbb{R}^{N \times K_N}$, where for $k = 1, \dots, K_N$, $x(k) = [x(k, 1), \dots, x(k, N)]^T \in \mathbb{R}^N$, we have

$$\|x\|_\infty^{(N)} = \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} |x(k, n)|.$$

Now we present a lemma on the relationship between the $z_N(k)$ and $f_N$.

*Lemma 4:* Assume that $z$ is continuously differentiable in $t$. Then for each $N$, there exists $u_N(k) \in \mathbb{R}^N$ such that for $k = 0, \dots, K_N - 1$,

$$z_N(k+1) - z_N(k) = \frac{1}{M} f_N(z_N(k)) + dt_N u_N(k), \tag{26}$$

and

$$\|u_N\|_\infty^{(N)} = O(\max\{\gamma_N, dt_N\}), \tag{27}$$

where $u_N = [u_N(0), \dots, u_N(K_N - 1)] \in \mathbb{R}^{N \times K_N}$.

*Proof:* Since $z$ is continuously differentiable in $t$, there exists $0 < c_1 < \infty$ such that for each $N$, for $k = 0, \dots, K_N - 1$ and $n = 1, \dots, N$, there exists a function $r_N : [0, T] \times \mathcal{D} \to \mathbb{R}$ such that

$$\begin{aligned} &\frac{z_N(k+1, n) - z_N(k, n)}{dt_N} \\ &= \frac{z(t_N(k), v_N(n)) - z(t_N(k), v_N(n))}{dt_N} \\ &= \dot{z}(t_N(k), v_N(n)) + r_N(t_N(k), v_N(n)), \end{aligned} \tag{28}$$

and for $N$ sufficiently large, $|r_N(t_N(k), v_N(n))| < c_1 dt_N$.

By (21) and (24), there exists $0 < c_2 < \infty$, such that for each $N$, for $k = 0, \dots, K_N - 1$ and $n = 1, \dots, N$, there exists a function $\phi_N : [0, T] \times \mathcal{D} \to \mathbb{R}$ such that

$$\begin{aligned} &\dot{z}(t_N(k), v_N(n)) \\ &= f(v_N(n), z_N(k, n), \nabla z_N(k, n), \nabla^2 z_N(k, n)) \\ &= f_N(z_N(k), v_N(n))/\delta_N + \phi_N(t_N(k), v_N(n)), \end{aligned} \tag{29}$$

and for $N$ sufficiently large, $|\phi_N(t_N(k), v_N(n))| < c_2 \gamma_N$, where $\{\gamma_N\}$ is as defined in (22).

For each $N$, for $k = 0, \dots, K_N - 1$ and $n = 1, \dots, N$, let

$$u_N(k, n) = \phi_N(t_N(k), v_N(n)) + r_N(t_N(k), v_N(n)),$$

and $u_N(k) = [u_N(k, 1), \dots, u_N(k, N)]^T \in \mathbb{R}^N$. Then there exists $0 < c < \infty$ such that for each $N$,

$$\|u_N\|_\infty^{(N)} < c \max\{\gamma_N, dt_N\}.$$

Hence (27) follows.

By (28) and (29), for each $N$, for $k = 0, \dots, K_N - 1$ and $n = 1, \dots, N$,

$$\frac{z_N(k+1) - z_N(k)}{dt_N} = \frac{f_N(z_N(k))}{\delta_N} + u_N(k).$$

By this and (25), we have (26). ■

In the following we show that under some conditions, $x_N(k)$ and $z_N(k)$ are asymptotically close for large $N$.

For each $N$, for $k = 0, \dots, K_N$ and $n = 1, \dots, N$, define

$$\varepsilon_N(k, n) = z_N(k, n) - x_N(k, n), \tag{30}$$

and let $\varepsilon_N(k) = [\varepsilon_N(k, 1), \dots, \varepsilon_N(k, N)]^T \in \mathbb{R}^N$.

By (18), (26), and (30), we have that for each $N$, for $k = 0, \dots, K_N$, there exists $u_N(k)$ as defined in Lemma 4 such that

$$\begin{aligned} \varepsilon_N(k+1) = \varepsilon_N(k) &+ \frac{1}{M}(f_N(z_N(k)) - f_N(x_N(k))) \\ &+ dt_N u_N(k). \end{aligned} \tag{31}$$

Suppose that for each $N$, $f_N \in \mathcal{C}^1$. Let $Df_N(x)$ be the derivative matrix of the function $f_N$ at $x$. Then we have that for each $N$, for $k = 1, \dots, K_N$ and $n = 1, \dots, N$, there exists a function $\tilde{f}_N : \mathbb{R}^N \to \mathbb{R}^N$ such that

$$f_N(z_N(k)) - f_N(x_N(k)) = Df_N(z_N(k))\varepsilon_N(k) + \tilde{f}_N(\varepsilon_N(k))$$

and
$$\tilde{f}_N(0) = 0. \tag{32}$$

Then we have from (31)
$$\begin{aligned}\varepsilon_N(k+1) = \varepsilon_N(k) + \frac{1}{M}(Df_N(z_N(k))\varepsilon_N(k) \\ + \tilde{f}_N(\varepsilon_N(k))) + dt_N u_N(k).\end{aligned} \tag{33}$$

Further suppose that for each $N$,
$$\|\varepsilon_N(0)\|_\infty^{(N)} = 0. \tag{34}$$

Define $\varepsilon_N = [\varepsilon_N(1), \ldots, \varepsilon_N(K_N)] \in \mathbb{R}^{N\times K_N}$. Then by (32), (33), and (34), for each $N$, there exists a function $H_N : \mathbb{R}^{N\times K_N} \to \mathbb{R}^{N\times K_N}$ such that
$$\varepsilon_N = H_N(u_N). \tag{35}$$

It follows that $H_N(0) = 0$ and $H_N \in \mathcal{C}^1$.

For each $N$, define
$$\mu_N = \lim_{\alpha\to 0} \sup_{\|u\|_\infty^{(N)}\le\alpha} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}}.$$

*Lemma 5:* Assume that

- $z$ is continuously differentiable in $t$;
- for each $N$, $f_N \in \mathcal{C}^1$;
- for each $N$, (34) holds; and
- the sequence $\{\mu_N\}$ is bounded.

Then
$$\|\varepsilon_N\|_\infty^{(N)} = O(\|u_N\|_\infty^{(N)}).$$

*Proof:* By definition, for each $N$, there exists $\delta > 0$ such that for $\alpha < \delta$,
$$\sup_{\|u\|_\infty^{(N)}\le\alpha} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}} < \mu_N + 1.$$

By (27), as $N \to \infty$, $\|u\|_\infty^{(N)} \to 0$. Then there exists $N_0$ and $\alpha_1$ such that for $N > N_0$, $\|u\|_\infty^{(N)} \le \alpha_1 < \delta$. Hence, for $N > N_0$,
$$\frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}} \le \sup_{\|u\|_\infty^{(N)}\le\alpha_1} \frac{\|H_N(u)\|_\infty^{(N)}}{\|u\|_\infty^{(N)}} < \mu_N + 1.$$

Therefore, there exists $0 < c < \infty$ such that for $N > N_0$,
$$\begin{aligned}\|\varepsilon_N\|_\infty^{(N)} = \|H_N(u_N)\|_\infty^{(N)} < (\mu_N + 1)\|u_N\|_\infty^{(N)} \\ < (c+1)\|u_N\|_\infty^{(N)}.\end{aligned}$$

This finishes the proof. ■

Lemma 5 states that as $N \to \infty$, $\|\varepsilon_N\|_\infty^{(N)} \to 0$, and at least with the same rate as $\|u_N\|_\infty^{(N)}$.

Let $X_N = [X_N(1)/M, \ldots, X_N(K_N)/M]$, $x_N = [x_N(1), \ldots, x_N(K_N)]$, and $z_N = [z_N(1), \ldots, z_N(K_N)]$, all in $\in \mathbb{R}^{N\times K_N}$. Now we present the main convergence theorem of this paper, which states that the value of the normalized Markov chain at time $k$ and node $n$, is close to that of $z$ at the corresponding point $(t_N(k), v_N(n)) \in [0,T]\times\mathcal{D}$ for large $M$ and $N$.

*Theorem 1:* Suppose that the assumptions in Lemma 1 and Lemma 5 hold. Then
$$\|X_N - z_N\|_\infty^{(N)} = O(\max\{\gamma_N, dt_N\}) \text{ a.s.}$$

*Proof:* By (27) and Lemma 5, there exists $0 < c_0 < \infty$ such that for $N$ sufficiently large,
$$\|\varepsilon_N\|_\infty^{(N)} < c_0 \max\{\gamma_N, dt_N\}. \tag{36}$$

Let $\tilde{T}_N$ in Corollary 1 be $T/\delta_N$. Then $\tilde{K}_N := \lfloor \tilde{T}_N M \rfloor = \lfloor T/dt_N \rfloor := K_N$. Hence by Corollary 1, for any sequence $\{\zeta_N\}$, for each $N$, we can take $M$ sufficiently large such that
$$\sum_{N=1}^{\infty} P\{\|X_N - x_N\|_\infty^{(N)} > \zeta_N\} \le \sum_{N=1}^{\infty} 1/N^2 < \infty.$$

By the first Borel-Cantelli Lemma [28],
$$P\left\{\lim\sup_{N\to\infty}\{\|X_N - x_N\|_\infty^{(N)} > \zeta_N\}\right\} = 0,$$

which implies that, a.s., for $N$ sufficiently large,
$$\|X_N - x_N\|_\infty^{(N)} < \zeta_N.$$

Take $\zeta_N$ such that for $N$ sufficiently large,
$$\zeta_N < \max\{\gamma_N, dt_N\}.$$

Then by the triangle inequality
$$\begin{aligned}\|X_N - z_N\|_\infty^{(N)} &\le \|X_N - x_N\|_\infty^{(N)} + \|x_N - z_N\|_\infty^{(N)} \\ &= \|X_N - x_N\|_\infty^{(N)} + \|\varepsilon_N\|_\infty^{(N)},\end{aligned}$$

a.s., there exists $0 < c < \infty$ such that for $N$ sufficiently large,
$$\|X_N - z_N\|_\infty^{(N)} \le c\max\{\gamma_N, dt_N\}.$$

This finishes the proof. ■

This theorem states that as $M \to \infty$ and $N \to \infty$, $X_N$ converges uniformly to $z_N$ a.s., and at least with the same rate as $\max\{\gamma_N, dt_N\}$.

### D. Convergence of Continuous-time-space Extension

In the following we study the convergence of the continuous-time-space extension of the Markov chain $X_N(k)$ to the PDE solution. Set $\tilde{T}_N = T/\delta_N$. For each $N$, we can construct $X_{oN}(\tilde{t})$ and $x_{oN}(\tilde{t})$ with time interval of length $1/M$, with $\tilde{t} \in [0, \tilde{T}_N]$. Respectively, let $X_{pN}(t)$ and $x_{pN}(t)$, where $t \in [0,T]$, be the continuous-space extension of $X_{oN}(\tilde{t})$ and $x_{oN}(\tilde{t})$ (with $\tilde{t} \in [0, \tilde{T}_N]$) by piecewise-constant space extensions on $\mathcal{D}$ and with time scaled by $\delta_N$ so that the time-interval length is $\delta_N/M := dt_N$. By ***piecewise-constant space extension*** of $X_{oN}$, we mean that we construct a piecewise-constant function on $\mathcal{D}$ such that the value of this function at each point in $\mathcal{D}$ is the value of the component of the vector $X_{oN}$ corresponding to the grid point that is "closest to the left" (taken one component at a time). Then for each $t$, $X_{pN}(t)$ and $x_{pN}(t)$ are real-valued functions defined on $\mathcal{D}$. Fig. 2 is an illustration of $x_N$ and $x_{pN}$ in a one-dimensional case. For fixed $T$, both $X_{pN}(t)$ and $x_{pN}(t)$ with $t \in [0,T]$ are in the space $D^{\mathcal{D}}[0,T]$ of functions of $[0,T]\times\mathcal{D} \to \mathbb{R}$ and are

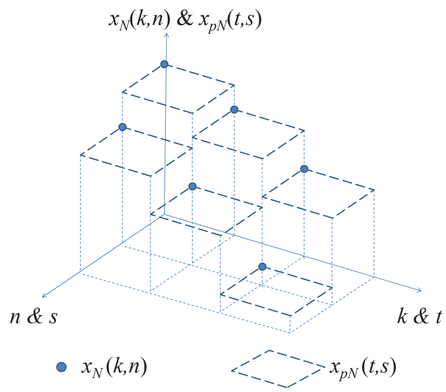


Fig. 2. An illustration of $x_N$ and $x_{pN}$ in a one-dimensional case.

Càdlàg with the time component. Define the $\infty$-norm $\|\cdot\|_\infty^{(p)}$ on $D^{\mathcal{D}}[0,T]$, i.e., for $x \in D^{\mathcal{D}}[0,T]$,

$$\|x\|_\infty^{(p)} = \sup_{\substack{t\in[0,T],\\ s\in\mathcal{D}}} |x(t,s)|.$$

First we show that $x_{pN}$ and $z$ are asymptotically close for large $N$.

*Lemma 6:* Suppose that the assumptions in Lemma 5 hold. Then

$$\|x_{pN} - z\|_\infty^{(p)} = O(\max\{\gamma_N, dt_N, ds_N\}).$$

*Proof:* For each $N$, for $k = 0, \dots, K_N$ and $n = 1, \dots, N$, by the definition of $x_{pN}$, we have that $x_{pN}(t_N(k), v_N(n)) = x_N(k,n)$. Let $\Omega_N(k,n)$ be the subset of $[0,T]\times\mathcal{D}$ containing $(t_N(k), v_N(n))$ where $x_{pN}$ is piecewise constant, i.e., $(t_N(k), v_N(n)) \in \Omega_N(k,n)$ and for all $(t,s) \in \Omega_N(k,n)$, $x_{pN}(t,s) = x_{pN}(t_N(k), v_N(n))$. (For example, for $\mathcal{D}\subset\mathbb{R}$, $\Omega_N(k,n) = [t_N(k), t_N(k+1)] \times [v_N(n), v_N(n+1)]$.) Then for each $N$,

$$\begin{aligned}\|x_{pN} - z\|_\infty^{(p)} &\le \|\varepsilon_N\|_\infty^{(N)} \\ &+ \max_{\substack{k=0,\dots,K_N\\ n=1,\dots,N}} \sup_{(t,s)\in\Omega_N(k,n)} |z(t_N(k), v_N(n)) - z(t,s)|.\end{aligned}$$

Since $z(t,s)$ is continuously differentiable in $t$ on a compact domain, it is Lipschitz continuous in $t$. Similarly, it is Lipschitz continuous in $s$. Hence there exist $0 < c_1, c_2 \le \infty$ such that for each $N$,

$$\begin{aligned}&\max_{\substack{k=0,\dots,K_N,\\ n=1,\dots,N}} \sup_{(t,s)\in\Omega_N(k,n)} |z(t_N(k), v_N(n)) - z(t,s)| \\ &\le c_1 \max_{\substack{k=0,\dots,K_N,\\ n=1,\dots,N}} \sup_{(t,s)\in\Omega_N(k,n)} \|(t_N(k), v_N(n)) - (t,s)\| \\ &\le c_2 \max\{ds_N, dt_N\},\end{aligned}$$

where $\|\cdot\|$ is some norm on $[0,T]\times\mathcal{D}$. Hence, by this and (36), there exists $0 < c < \infty$ such that for $N$ sufficiently large,

$$\|x_{pN} - z\|_\infty^{(p)} \le c\max\{\gamma_N, dt_N, ds_N\}.$$

This finishes the proof. ■

Now we present a convergence theorem for the continuous functions.

*Theorem 2:* Suppose that the assumptions in Lemma 1 and Lemma 5 hold. Then

$$\|X_{pN} - z\|_\infty^{(p)} = O(\max\{\gamma_N, dt_N, ds_N\}) \text{ a.s. on } [0,T]\times\mathcal{D}.$$

*Proof:* By Lemma 3 , for any sequence $\{\zeta_N\}$, for each $N$, we can take $M$ sufficiently large such that

$$\sum_{N=1}^{\infty} P\{\|X_{oN} - x_{oN}\|_\infty^{(o)} > \zeta_N\} \le \sum_{N=1}^{\infty} 1/N^2 < \infty.$$

By the first Borel-Cantelli Lemma [28],

$$P\left\{\lim_{N\to\infty}\sup \{\|X_{oN} - x_{oN}\|_\infty^{(o)} > \zeta_N\ \}\right\} = 0,$$

which implies that, a.s., for $N$ sufficiently large,

$$\|X_{oN} - x_{oN}\|_\infty^{(o)} < \zeta_N \text{ on } [0, \tilde{T}_N].$$

Since $X_{pN}$ and $x_{pN}$ are the piecewise continuous-space extensions of $X_{oN}$ and $x_{oN}$ by constant interpolation, respectively, it follows that for any sequence $\{\zeta_N\}$, we can take $M$ sufficiently large such that, a.s., for $N$ sufficiently large,

$$\|X_{pN} - x_{pN}\|_\infty^{(p)} < \zeta_N \text{ on } [0,T]\times\mathcal{D}.$$

Take $\zeta_N$ such that for $N$ sufficiently large,

$$\zeta_N < \max\{\gamma_N, dt_N, ds_N\}.$$

Then by the triangle inequality

$$\|X_{pN} - z\|_\infty^{(p)} \le \|X_{pN} - x_{pN}\|_\infty^{(p)} + \|x_{pN} - z\|_\infty^{(p)}$$

and Lemma 6, a.s., there exists $0 < c < \infty$ such that for $N$ sufficiently large,

$$\|X_{pN} - z\|_\infty^{(p)} \le c\max\{\gamma_N, dt_N, ds_N\} \text{ on } [0,T]\times\mathcal{D}.$$

This finishes the proof. ■

This theorem states that as $M \to \infty$ and $N \to \infty$, the continuous-time-space extension $X_{pN}$ of the Markov chain $X_N(k)$, converges uniformly to $z$, the solution of the PDE a.s., and at least with the same rate as $\max\{\gamma_N, dt_N, ds_N\}$.

The solution of the PDE can be found quickly by mathematical tools readily available and then be used to approximate the Markov chain $X_N(k)$. We give an example of this in the next section.

## IV. Application to the Modeling of Large Networks

In this section we present an example of the application of our approach to network modeling. We show how the Markov chain representing the queue lengths of the nodes in the network can be approximated by the solution of a PDE using the results of the preceding section.

$$
X_N(k+1,n) - X_N(k,n) = \begin{cases} 1+G(k,n), \text{ with probability} \\ (1-W(n,X_N(k,n)/M)) \\ \times [P_r(n-1)W(n-1,X_N(k,n-1)/M)(1-W(n+1,X_N(k,n+1)/M)) \\ + P_l(n+1)W(n+1,X_N(k,n+1)/M)(1-W(n-1,X_N(k,n-1)/M))]; \\ -1+G(k,n), \text{ with probability} \\ W(n,X_N(k,n)/M) \\ \times [P_r(n)(1-W(n+1,X_N(k,n+1)/M))(1-W(n+2,X_N(k,n+2)/M)) \\ + P_l(n)(1-W(n-1,X_N(k,n-1)/M))(1-W(n-2,X_N(k,n-2)/M))]; \\ G(k,n), \text{ otherwise.} \end{cases} \tag{37}
$$

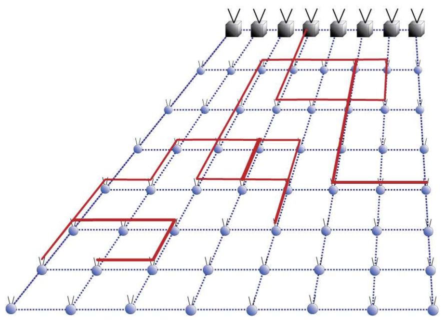

Fig. 3. An illustration of a wireless sensor network over a two-dimensional domain. Destination nodes are located at the far edge. We show the possible path of a message originating from a node located in the left-front region.

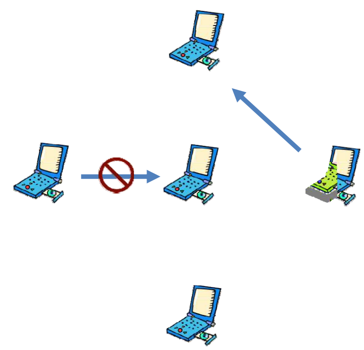

Fig. 4. An illustration of the collision protocol: reception at a node fails when more than one of its neighbors transmit (regardless of the intended receiver).

### A. Network Model

We consider a network of wireless sensor nodes uniformly placed over a domain. In a random fashion, the sensor nodes generate data messages that need to be communicated to the destination nodes located on the boundary of the domain, which represent specialized devices that collect the sensor data. The sensor nodes also serve as relays in the routing of the messages to the destination nodes. Each sensor node has the capacity to store messages and decides to transmit or receive messages to or from its immediate neighbors at each time instant, but not both. This simplified rule of transmission allows for a relatively simple representation. We illustrate such a network over a two-dimensional domain in Fig. 3. The communication is interference-limited because all nodes share the same wireless channel. We assume a simple collision protocol: a transmission from a transmitter to a neighboring receiver is successful if and only if none of the other neighbors of the receiver is a transmitter, as illustrated in Fig. 4.

### B. Continuum Model in One Dimension

We first consider the case of a one-dimensional network, where $N$ sensor nodes are uniformly placed over a domain $\mathcal{D} \subset \mathbb{R}$ and labeled by $n = 1, \ldots, N$. The destination nodes are located on the boundary of $\mathcal{D}$, labeled $n = 0$ and $n = N+1$. Again let $ds_N$ be the distance between neighboring nodes. Let $X_N(k,n)$ in (15) be the queue length of node $n$ at time $k$. Let $M$ in (16) be the maximum queue length of each node.

At each time instant $k = 0, 1, \ldots$, node $n$ decides to be a transmitter with probability $W(n, X_N(k,n)/M)$. Assume that node $n$ randomly chooses to transmit to the right or the left immediate neighbor with probability $P_r(n)$ and $P_l(n)$, respectively. Define $G(k) = [G(k,1), \ldots, G(k,N)]^T$, where $G(k,n)$ is the number of messages generated at node $n$ at time $k$. We model $G(k,n)$ by independent Poisson random variables with mean $g(n)$. The destination nodes at the boundaries of the domain do not have queues; they simply receive any message transmitted to it and never itself transmits anything. We illustrate the time evolution of the queues in the network in Fig. 5.

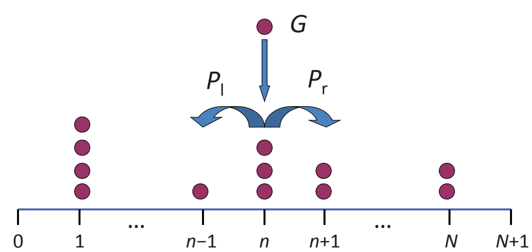


Fig. 5. An illustration of the time evolution of the queues in the one-dimensional network model.

The sequence $X_N(k)$ defined above forms a Markov chain whose evolution is described by (16). According to the behavior of the nodes, the $n$th component of $F_N(X_N(k)/M, U(k))$, where $n = 1, \ldots, N$, is defined by (37) at the top of the page, where $X_N(k,n)$ with $n \le 0$ or $n \ge N+1$ are defined to be zero. For simplicity, in the following parts we set $W(n, X_N(k,n)/M) = X_N(k,n)/M$, which corresponds to the transmission rule that a node transmits a message with a probability proportional to its queue length. With this

simplification, for $x = [x_1, \dots, x_N]^T$, the $n$th component of $F_N(x, U(k))$, where $n = 1, \dots, N$, is

$$\begin{cases} 1 + G(k,n), \text{ with probability} \\ (1-x_n)[P_r(n-1)x_{n-1}(1-x_{n+1}) \\ \quad + P_l(n+1)x_{n+1}(1-x_{n-1})]; \\ -1 + G(k,n), \text{ with probability} \\ x_n[P_r(n)(1-x_{n+1})(1-x_{n+2}) \\ \quad + P_l(n)(1-x_{n-1})(1-x_{n-2})]; \\ G(k,n), \text{ otherwise,} \end{cases}$$

where $x_n$ with $n \le 0$ or $n \ge N+1$ are defined to be zero.

Define $f_N$ as in (17). It follows that for $x = [x_1, \dots, x_N]^T$, the $n$th component of $f_N(x)$, where $n = 1, \dots, N$, is

$$\begin{aligned} &(1-x_n)[P_r(n-1)x_{n-1}(1-x_{n+1}) \\ &+ P_l(n+1)x_{n+1}(1-x_{n-1})] \\ &- x_n[P_r(n)(1-x_{n+1})(1-x_{n+2}) \\ &+ P_l(n)(1-x_{n-1})(1-x_{n-2})] + g(n), \end{aligned} \tag{38}$$

where $x_n$ with $n \le 0$ or $n \ge N+1$ are defined to be zero. Define the deterministic sequence $x_N(k)$ as in (18).

Set $\delta_N$, defined in Section III-C, to be $ds_N^2$. Let

$$dt_N = \delta_N / M = ds_N^2 / M. \tag{39}$$

Assume

$$P_l(n) = p_l(v_N(n)) \text{ and } P_r(n) = p_r(v_N(n)), \tag{40}$$

where $p_l(s)$ and $p_r(s)$ are real-valued functions defined on $\mathcal{D}$. As in Section II we again assume

$$p_l(s) = b(s) + c_l(s)ds_N \text{ and } p_r(s) = b(s) + c_r(s)ds_N. \tag{41}$$

Let $c = c_l - c_r$. Again we call $b$ the diffusion and $c$ the convection. In order to guarantee that the number of messages entering the system from outside over finite time intervals remains finite throughout the limiting process, we set $g(n) = Mg_p(v_N(n))dt_N$. Assume $b, c_l, c_r$, and $g_p$ are in $\mathcal{C}^1$. Then $f_N \in \mathcal{C}^1$.

Let $f_N(y_N, s_N)$ be defined as in Section III-C. Then we have the $f$ in (21):

$$\begin{aligned} f = {}& b(s)\frac{d}{ds}\left((1-z(s))(1+3z(s))z_s(s)\right) \\ &+ 2(1-z(s))z_s(s)b_s(s) + z(s)(1-z(s))^2 b_{ss}(s) \\ &+ \frac{d}{ds}(c(s)z(s)(1-z(s))^2) + g_p(s). \end{aligned} \tag{42}$$

Here, recall that, a single subscript $s$ represents first derivative and a double subscript $ss$ represents second derivative.

Based on the behavior of nodes $n = 1$ and $n = N$ next to the destination nodes, we derive the boundary condition for the PDE. For example, the node $n = 1$ receives messages only from the right and encounters no interference when transmitting to the left. Replacing $x_n$ with $n \le 0$ or $n \ge N+1$ by 0 in (38), it follows that the 1st component of $f_N(x)$ is

$$\begin{aligned} &(1-x_n)P_l(n+1)x_{n+1} \\ &- x_n[P_l(n) + P_r(n)(1-x_{n+1})(1-x_{n+2})] + g(n). \end{aligned} \tag{43}$$

Similarly, the $N$th component of $f_N(x)$ is

$$\begin{aligned} &(1-x_n)P_r(n-1)x_{n-1} \\ &- x_n[P_r(n) + P_l(n)(1-x_{n-1})(1-x_{n-2})] + g(n). \end{aligned} \tag{44}$$

Set $\beta_N$, defined in Section III-C, to be 1. Then we have the $h$ in (23):

$$h = -b(s)z(s)^3 + b(s)z(s)^2 - b(s)z(s). \tag{45}$$

Solving $h = 0$ for real $z$, we have the boundary condition $z(t,s) = 0$. This equation might seem confusing to some readers as the limit of (43) and (44), if it has not been noticed that, unlike $f$, $g$ is the limit of a different function $f_N(y_N, s_N)/\beta_N$.

For fixed $T$, let $z : [0,T] \times \mathcal{D} \to \mathbb{R}$ be the solution of the PDE (24), with boundary condition $z(t,s) = 0$ and initial condition $z(0,s) = z_0(s)$, where the right hand side of (24) is

$$\begin{aligned} & b(s)\frac{\partial}{\partial s}\left((1-z(t,s))(1+3z(t,s))z_s(t,s)\right) \\ &+ 2(1-z(t,s))z_s(t,s)b_s(s) + z(t,s)(1-z(t,s))^2 b_{ss}(s) \\ &+ \frac{\partial}{\partial s}(c(s)z(t,s)(1-z(t,s))^2) + g_p(s). \end{aligned} \tag{46}$$

In the following we show the convergence of the Markov chain $X_N(k)$ to the PDE solution $z$ to in the one-dimensional network case. Define $K_N$, $z_N$, $u_N$, and $\varepsilon_N$ as in Section III-C. Throughout this section we assume (34) holds. By (38) and (46), it follows that there exists $0 < c < \infty$ such that for $N$ sufficiently large,

$$\|\gamma_N\|_\infty^{(N)} < c ds_N. \tag{47}$$

Albeit arduous, the algebraic manipulation in getting (42), (45), and (47) amounts only to algebraic exercises, the concept of which is no more sophisticated than that in getting (14) in Section II. In practice, we accomplishe such manipulation using symbolic tools provided by computer programs such as Matlab.

By (35), for each $N$, for $k = 1, \dots, K_N$ and $n = 1, \dots, N$, we can write $\varepsilon_N(k,n) = H_N^{(k,n)}(u_N)$, where $H_N^{(k,n)}$ is a real-valued function defined on $\mathbb{R}^{N \times K_N}$. It follows that $H_N^{(k,n)}(0) = 0$ and $H_N^{(k,n)} \in \mathcal{C}^1$.

Define

$$DH_N = \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \left| \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) \right|,$$

where 0 is in $\mathbb{R}^{N \times K_N}$.

*Lemma 7:* We have that for each $N$,

$$\mu_N \le DH_N.$$

*Proof:* For each $N$, we have

$$\begin{aligned} &\max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left| \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right| \\ &\le \max_{\substack{k=1,\dots,K_N \\ n=1,\dots,N}} \left( \sum_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} \left| \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) \right| |u(i,j)| \right) \\ &\le DH_N \max_{\substack{i=1,\dots,K_N \\ j=1,\dots,N}} |u(i,j)| \end{aligned}$$

$$= DH_N \|u\|_\infty^{(N)}.$$

Thus, for each $N$, for all $u \neq 0$,

$$DH_N \geq \frac{\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right|}{\|u\|_\infty^{(N)}}. \tag{48}$$

For each $N$, let $v = [v(1), \ldots, v(K_N)]$, where $v(k) = [v(k,1), \ldots, v(k,N)]^T$, where for $k = 1, \ldots, K_N$ and $n = 1, \ldots, N$,

$$v(k,n) = \operatorname{sgn} \frac{\partial H_N^{(k_0,n_0)}}{\partial u(k,n)}(0),$$

where

$$(k_0, n_0) \in \arg\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \left| \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) \right|.$$

Then

$$DH_N = \frac{\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) v(i,j) \right|}{\|v\|_\infty^{(N)}}.$$

By this and (48) we have

$$DH_N = \sup_{u \neq 0} \frac{\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right|}{\|u\|_\infty^{(N)}}. \tag{49}$$

By Taylor's theorem, for each $N$, for $k = 1, \ldots, K_N$ and $n = 1, \ldots, N$, there exists $\tilde{H}_N^{(k,n)}(u)$ such that

$$H_N^{(k,n)}(u) = \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) + \tilde{H}_N^{(k,n)}(u), \tag{50}$$

and for $i = 1, \ldots, K_N$ and $j = 1, \ldots, N$,

$$\lim_{u \to 0} \frac{|\tilde{H}_N^{(k,n)}(u)|}{\|u\|_\infty^{(N)}} = 0.$$

Hence for each $\varepsilon > 0$, there exists $\delta$ such that for $\|u\|_\infty^{(N)} < \delta$, we have

$$\frac{|\tilde{H}_N^{(k,n)}(u)|}{\|u\|_\infty^{(N)}} < \varepsilon.$$

Then for $\|u\|_\infty^{(N)} \leq \alpha \leq \delta$,

$$\sup_{\|u\|_\infty^{(N)} \leq \alpha} \frac{|\tilde{H}_N^{(k,n)}(u)|}{\|u\|_\infty^{(N)}} < \varepsilon.$$

Therefore, for $i = 1, \ldots, K_N$ and $j = 1, \ldots, N$,

$$\lim_{\alpha \to 0} \sup_{\|u\|_\infty^{(N)} \leq \alpha} \frac{|\tilde{H}_N^{(k,n)}(u)|}{\|u\|_\infty^{(N)}} = 0. \tag{51}$$

By (50), for each $N$,

$$\|H_N(u)\|_\infty^{(N)} \leq \max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \tilde{H}_N^{(k,n)}(u) \right|$$
$$+ \max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \partial H_N^{(k,n)} \partial u(i,j)(0) u(i,j) \right|.$$

Hence

$$\mu_N \leq \lim_{\alpha \to 0} \sup_{\|u\|_\infty^{(N)} \leq \alpha} \left( \frac{\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \tilde{H}_N^{(k,n)}(u) \right|}{\|u\|_\infty^{(N)}} + \frac{\max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \left| \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \frac{\partial H_N^{(k,n)}}{\partial u(i,j)}(0) u(i,j) \right|}{\|u\|_\infty^{(N)}} \right).$$

Hence by (49) and (51), we finish the proof. ∎

Notice that $DH_N$ is essentially the induced $\infty$-norm of the linearized version of the operator $H_N$.

Now we present a lemma on the condition of the sequence $\{\mu_N\}$ being bounded for the one-dimensional network case.

*Lemma 8:* In the one-dimensional network case, assume that the function

$$\max\{|z|, |z_s|, |z_{ss}|, |b|, |b_s|, |b_{ss}|, |c|, |c_s|\} \tag{52}$$

of $(t, s)$ is bounded on $[0, T] \times \mathcal{D}$. Then $\{\mu_N\}$ is bounded.

*Proof:* Define

$$A_N(k) = I_N + \frac{1}{M} Df_N(z_N(k)),$$

where $I_N$ be the identity matrix in $\mathbb{R}^{N \times N}$. It follows from (33) that for each $N$ and for $k = 0, \ldots, K_N$,

$$\varepsilon_N(k+1) = A_N(k)\varepsilon_N(k) + \frac{\tilde{f}_N(\varepsilon_N(k))}{M} + dt_N u_N(k).$$

It follows that

$$\begin{aligned}
\varepsilon_N(k) = dt_N(&A_N(k-1)\ldots A(1)u_N(0)\\
&+ A_N(k-1)\ldots A(2)u_N(1)\\
&+ \ldots\\
&+ A_N(k-1)u_N(k-2) + u_N(k-1))\\
&+ \frac{1}{M}(A_N(k-1)\ldots A(2)\tilde{f}_N(\varepsilon_N(1))\\
&+ A_N(k-1)\ldots A(3)\tilde{f}_N(\varepsilon_N(2))\\
&+ \ldots\\
&+ A_N(k-1)\tilde{f}_N(\varepsilon_N(k-2))\\
&+ \tilde{f}_N(\varepsilon_N(k-1)).
\end{aligned}$$

Define

$$B_N^{(k,n)} = \begin{cases} 0, & 0 \leq n < k-3;\\ I_N, & n = k-3;\\ A_N(k-1)\ldots A_N(n+1), & n \geq k-2. \end{cases} \tag{53}$$

It follows that

$$\frac{\partial H_N^{(k,n)}(u)}{\partial u(i,j)}(0) = B_N^{(k,i)}(n,j) dt_N.$$

Hence by Lemma 7,

$$\mu_N \leq \max_{\substack{k=1,\ldots,K_N\\ n=1,\ldots,N}} \sum_{\substack{i=1,\ldots,K_N\\ j=1,\ldots,N}} \left| B_N^{(k,i)}(n,j) \right| dt_N. \tag{54}$$

By (38), for fixed $N$, for $x = [x_1, \dots, x_N]^T$, the $(n,m)$th component of $Df_N(x) := \frac{\partial f_N^{(n)}}{\partial x_m}(x)$, where $n, m = 1, \dots, N$, is

$$\begin{cases} P_l(n)x_n(1-x_{n-1}), & m = n-2; \\ (1-x_n)[P_r(n-1)(1-x_{n+1}) \\ \quad -P_l(n+1)x_{n+1}] \\ \quad +P_l(n)x_n(1-x_{n-2}), & m = n-1; \\ -[P_r(n-1)x_{n-1}(1-x_{n+1}) \\ \quad +P_l(n+1)x_{n+1}(1-x_{n-1})] \\ \quad -[P_r(n)(1-x_{n+1})(1-x_{n+2}) \\ \quad +P_l(n)(1-x_{n-1})(1-x_{n-2})], & m = n; \\ (1-x_n)[P_l(n+1)(1-x_{n-1}) \\ \quad -P_r(n-1)x_{n-1}] \\ \quad +P_r(n)x_n(1-x_{n+2})], & m = n+1; \\ P_r(n)x_n(1-x_{n+1}), & m = n+2; \\ 0 & \text{other wise,} \end{cases}$$

where $x_n$ with $n \le 0$ or $n \ge N+1$ are defined to be zero.

Denote the induced $\infty$-norm on $\mathbb{R}^{N\times N}$ again by $\|\cdot\|_\infty^{(N)}$. That is, for $A \in \mathbb{R}^{N\times N}$, with the $(i,j)$th element being $A(i,j)$,

$$\|A\|_\infty^{(N)} = \max_{1\le i\le N} \sum_{j=1}^{N} |A(i,j)|,$$

which is simply the maximum absolute row sum of the matrix. Then we have,

$$\begin{aligned} &\|A_N(k)\|_\infty^{(N)} \\ &= \max_{n=1,\dots,N} \frac{1}{M}(|P_l(n)z_N(k,n)(1-z_N(k,n-1))| \\ &\quad + |(1-z_N(k,n))[P_r(n-1)(1-z_N(k,n+1)) \\ &\quad - P_l(n+1)z_N(k,n+1)] \\ &\quad + P_l(n)z_N(k,n)(1-z_N(k,n-2))| \\ &\quad + |M - [P_r(n-1)z_N(k,n-1)(1-z_N(k,n+1)) \\ &\quad + P_l(n+1)z_N(k,n+1)(1-z_N(k,n-1))] \\ &\quad - [P_r(n)(1-z_N(k,n+1))(1-z_N(k,n+2)) \\ &\quad + P_l(n)(1-z_N(k,n-1))(1-z_N(k,n-2))]| \\ &\quad + |(1-z_N(k,n))[P_l(n+1)(1-z_N(k,n-1)) \\ &\quad - P_r(n-1)z_N(k,n-1)] \\ &\quad + P_r(n)z_N(k,n)(1-z_N(k,n+2))]| \\ &\quad + |P_r(n)z_N(k,n)(1-z_N(k,n+1))|). \end{aligned}$$

Put (40), (41), and the Taylor's expansions (11), (12), and (13) of $z, b$, and $c$, respectively, into the above equation and rearrange. (Again we omit the detailed algebraic manipulation here.) Then we have that there exist $0 < c_1 < \infty$ such that for each $N$, for $k = 1, \dots, K_N$ and $n = 1, \dots, N$,

$$\begin{aligned} &\|A_N(k)\|_\infty^{(N)} \\ &\le \max_{n=1,\dots,N} |-c_s(v_N(n)) - b_{ss}(v_N(n)) \\ &\quad - 2b(v_N(n))z_{ss}(t_N(k),v_N(n)) \\ &\quad + 4b_{ss}(v_N(n))z(t_N(k),v_N(n)) \\ &\quad + 2b_s(v_N(n))z_s(t_N(k),v_N(n)) \\ &\quad + 4c_s(v_N(n))z(t_N(k),v_N(n)) \\ &\quad + 4c(v_N(n))z_s(t_N(k),v_N(n)) \\ &\quad + 6b(v_N(n))z_s(t_N(k),v_N(n))^2 \\ &\quad - 3b_{ss}(v_N(n))z(t_N(k),v_N(n))^2 \\ &\quad - 3c_s(v_N(n))z(t_N(k),v_N(n))^2 \\ &\quad + 6b(v_N(n))z(t_N(k),v_N(n))z_{ss}(t_N(k),v_N(n)) \\ &\quad - 6c(v_N(n))z(t_N(k),v_N(n))z_s(t_N(k),v_N(n))|\frac{ds^2}{M} \\ &\quad + c_1\frac{ds^3}{M} + 1 \\ &:= \max_{n=1,\dots,N} |q(t_N(k),v_N(n))|\frac{ds^2}{M} + c_1\frac{ds^3}{M} + 1. \end{aligned}$$

Since (52) is bounded, there exists $0 < c_2 < \infty$ such that $|q(t,s)| < c_2$ for all $(t,s) \in [0,T]\times\mathcal{D}$. Hence for each $N$ and for $k = 0, \dots, K_N$,

$$\|A_N(k)\|_\infty^{(N)} \le 1 + c_2\frac{ds_N^2}{M} + c_1\frac{ds_N^3}{M}.$$

Hence there exists $0 < c_3 < \infty$, for $N$ sufficiently large and for $k = 0, \dots, K_N$,

$$\|A_N(k)\|_\infty^{(N)} \le 1 + c_3\frac{ds_N^2}{M} = 1 + c_3 dt_N.$$

Hence by (53) and (54), for $N$ sufficiently large,

$$\mu_N \le K_N dt_N (1 + c_3 dt_N)^{K_N}.$$

Since $T < \infty$, there exist $0 < c_4 < \infty$ such that for each $N$, $K_N dt_N < c_4$. But as $N \to \infty$, $K_N \to \infty$, and

$$(1 + c_3 dt_N)^{K_N} = \left(1 + \frac{c_3 T}{K_N}\right)^{K_N} \to e^{c_3 T}.$$

Therefore, there exist $0 < c_5 < \infty$ such that for each $N$, $\mu_N < c_5$. This finishes the proof. ■

*Proposition 1:* In the one-dimensional network case, suppose that the assumption in Lemma 8 holds. Then

$$\|X_N - z_N\|_\infty^{(N)} = O(ds_N) \text{ a.s. on } [0,T]\times\mathcal{D}.$$

*Proof:* By (39) and (47), there exists $0 < c < \infty$ such that for $N$ sufficiently large,

$$\max\{\gamma_N, ds_N, dt_N\} \le c\, ds_N.$$

One can now easily verify that the assumptions in Theorem 1 hold. Then by Theorem 1 the desired result holds. ■

This proposition states that in the one-dimensional network case, as $M \to \infty$ and $N \to \infty$, $X_N$ converges uniformly to $z_N$ a.s., and at least with the same rate as $ds_N$. Analogously, for the continuous-time-space extension $X_{pN}$ of $X_N(k)$, given the same assumption as in the above theorem, by Theorem 2, we have

$$\|X_{pN} - z\|_\infty^{(p)} = O(ds_N) \text{ a.s. on } [0,T]\times\mathcal{D}.$$

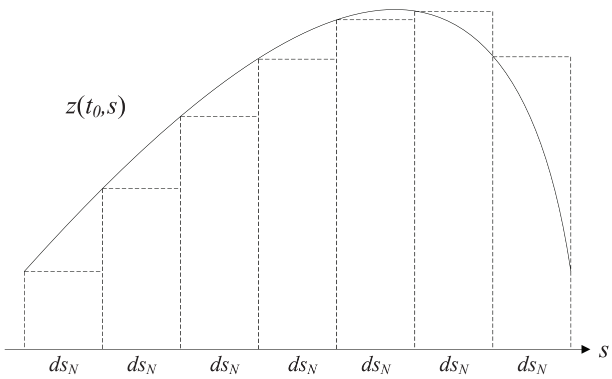


Fig. 6. The PDE solution $z(t,s)$, at $t = t_o$ approximating the normalized queue lengths of a one-dimensional network.

*1) Interpretation of the approximation PDE:* Now we make some remarks on how to use a given approximating PDE. First, for fixed $N$ and $M$, the normalized queue length of node $n$ at time $k$, is approximated by the value of the PDE solution $z$ at the corresponding point in $[0,T] \times \mathcal{D}$, i.e.,

$$z((t_N(k), v_N(n))) \approx \frac{X_N(k,n)}{M}.$$

Second, we show how to interpret

$$C(t_o) := \int_{\mathcal{D}} z(t_o,s)ds_N,$$

the area below the curve $z(t_o,s)$ for fixed $t_o \in [0,T]$. Let $k_o = \lfloor t_o/dt_N \rfloor$. Then we have

$$z(t_o, v_N(n))ds_N \approx \frac{X_N(k_o,n)}{M} ds_N,$$

the area of the $n$th rectangle in Fig. 6. Hence

$$C(t_o) \approx \sum_{n=1}^{N} z(t_o, v_N(n))ds_N \approx \sum_{n=1}^{N} \frac{X_N(k_o,n)}{M} ds_N,$$

the sum of all rectangles. If we assume that all messages in the queue have roughly the same bits, and think of $ds_N$ as the "coverage" of each node, then the area under any segment of the curve measures a kind of "data-coverage product" of the nodes covered by the segment, in the unit of "bit·meter". As $N \to \infty$, the total normalized queue length $\sum_{n=1}^{N} X_N(k_o,n)/M$ of the network does go to infinity; however, the coverage $ds_N$ of each node goes to 0. Hence the sum of the "data-coverage product" can be approximated by the finite area $C(t_o)$.

*2) Comparison between PDE approximation and Monte Carlo simulation: One dimension:* We compare the PDE approximation obtained from our approach with the Monte Carlo simulations for a network over the domain $\mathcal{D} = [-1,1]$. We use the initial condition $z_0(s) = l_1 e^{-s^2}$, where $l_1 > 0$ is a constant, so that initially the nodes in the middle have messages to transmit, while those near the boundaries have very few. We set the message generation rate $g_p(s) = l_2 e^{-s^2}$, where $l_2 > 0$ is a parameter determining the total load of the system.

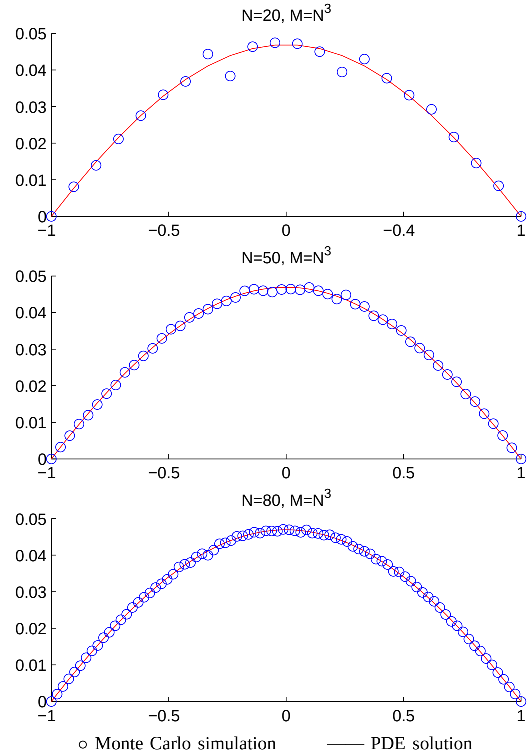


Fig. 7. The Monte Carlo simulations (with different $N$ and $M$) and the PDE solution of a one-dimensional network, with $b = 1/2$ and $c = 0$, at $t = 1s$.

We use three sets of values of $N = 20, 50, 80$ and $M = N^3$, and show the PDE solution and the Monte Carlo simulation results with different $N$ and $M$ at $t = 1s$. The networks have diffusion coefficient $b = 1/2$ and convection coefficient $c = 0$ in Fig. 7 and $c = 1$ in Fig. 8, respectively, where the x-axis denotes the node location and y-axis denotes the normalized queue length.

For the three sets of the values of $N = 20, 50, 80$ and $M = N^3$ and with $c = 0$, the maximum absolute errors of the PDE approximation are $5.6 \times 10^{-3}$, $1.3 \times 10^{-3}$, and $1.1 \times 10^{-3}$, respectively; and with $c = 1$, the errors are $4.4 \times 10^{-3}$, $1.5 \times 10^{-3}$, and $1.1 \times 10^{-3}$, respectively. As we can see, as $N$ and $M$ increase, the resemblance between the Monte Carlo simulations and the PDE solution becomes stronger. In the case of very large $N$ and $M$, it is difficult to distinguish the results.

We stress that the PDEs only took fractions of a second to solve on a computer, while the Monte Carlo simulations took time on the order of tens of hours. We could not do Monte Carlo simulations of any larger networks because of prohibitively long computation time.

## C. Continuum Model in Two Dimensions

Generalization of the continuum model to higher dimensions is straightforward, except for more arduous algebraic manipulation. Now we consider the two-dimensional network of $N_1 \times N_2$ sensor nodes. The nodes are uniformly placed over the domain $\mathcal{D} \subset \mathbb{R}^2$ and labeled by $(n,m)$, where $n = 1, \ldots, N_1$ and $m = 1, \ldots, N_2$. Again let the distance between neighboring nodes be $ds_N$. Assume that the node at

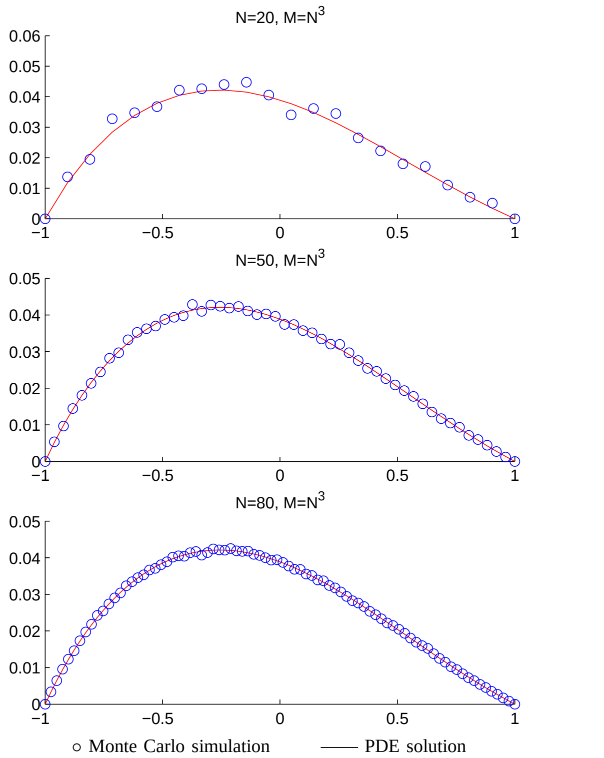


Fig. 8. The Monte Carlo simulations (with different $N$ and $M$) and the PDE solution of a one-dimensional network, with $b = 1/2$ and $c = 1$, at $t = 1s$.

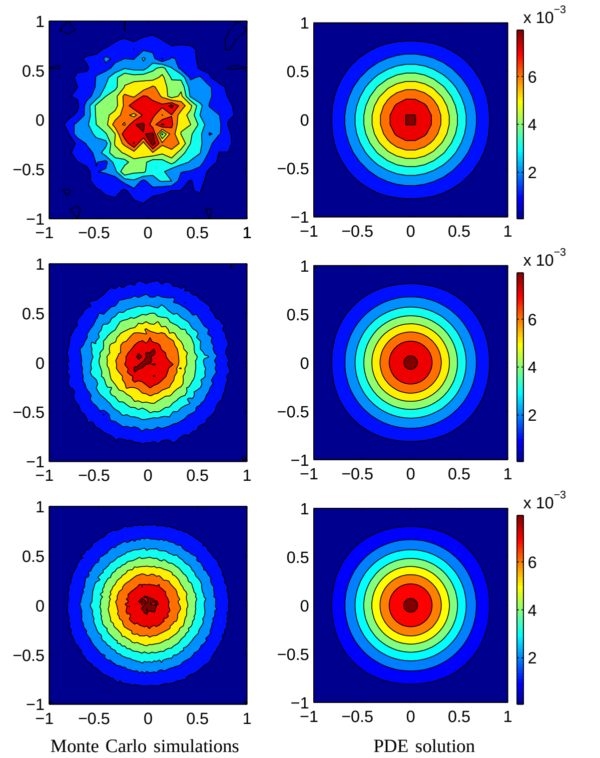


Fig. 9. The Monte Carlo simulations (from top to bottom, with $N_1 = N_2 = 20, 50, 80$, respectively, and $M = N_1^3$) and the PDE solution of a two-dimensional network, with $b_1 = b_2 = 1/4$ and $c_1 = c_2 = 0$, at $t = 0.1s$.

location $(n, m)$ randomly chooses to transmit to the north, east, south, or west immediate neighbor with probabilities $P_e(n,m) = b_1(s) + c_e(s)ds_N$, $P_w(n,m) = b_1(s) + c_w(s)ds_N$, $P_n(n,m) = b_2(s) + c_n(s)ds_N$, and $P_s(n,m) = b_2(s) + c_s(s)ds_N$, respectively. Define $c_1 = c_w - c_e$ and $c_2 = c_s - c_n$.

The derivation of the approximating PDE is similar to those of the one-dimensional cases, except that we now have to consider transmission to and interference from four directions instead of two. We present the approximating PDE here without the detailed derivation:

$$\dot{z} = \sum_{j=1}^{2} b_j \frac{\partial}{\partial s_j}\left((1+5z)(1-z)^3 \frac{\partial z}{\partial s_j}\right) + 2(1-z)^3 \frac{\partial z}{\partial s_j} \times \frac{db_j}{ds_j} + z(1-z)^4 \frac{d^2 b_j}{ds_j^2} + \frac{\partial}{\partial s_j}\left(c_j z(1-z)^4\right),$$

with boundary condition $z(t, s) = 0$ and initial condition $z(0, s) = z_0(s)$, where $t \in [0, T]$ and $s = (s_1, s_2) \in \mathcal{D}$.

*1) Comparison between PDE approximation and Monte Carlo simulations: Two dimensions:* We compare the PDE approximation and the Monte Carlo simulations of a network over the domain $D = [-1, 1] \times [-1, 1]$. We use the initial condition $z_0(s) = l_1 e^{-(s_1^2 + s_2^2)}$, where $l_1 > 0$ is a constant, so that initially the nodes in the center have messages to transmit, while those near the boundary have very few. We set the message generation rate $g_p(s) = l_2 e^{-(s_1^2+s_2^2)}$, where $l_2 > 0$ is a parameter determining the total load of the system.

We use three different sets of the values of $N_1 \times N_2$ and $M$, where $N_1 = N_2 = 20, 50, 80$ and $M = N_1^3$. We show the contours of the normalized queue length from the PDE solution and the Monte Carlo simulation results with different sets of values of $N_1$, $N_2$, and $M$, at $t = 0.1s$. The networks have diffusion coefficients $b_1 = b_2 = 1/4$ and convection coefficients $c_1 = c_2 = 0$ and $c_1 = -2, c_2 = -4$ in Fig. 9 and Fig. 10, respectively. It took 3 days to do the Monte Carlo simulation of the network at $t = 0.1s$ with $80 \times 80$ nodes and the maximum queue length $M = 80^3$, while the PDE solved on the same machines took less than a second. We could not do Monte Carlo simulations of any larger networks or greater values of $t$.

For the three sets of values of $N_1 = N_2 = 20, 50, 80$ and $M = N_1^3$ and with $c_1 = c_2 = 0$, the maximum absolute errors are $3.2 \times 10^{-3}$, $1.1 \times 10^{-3}$, and $6.8 \times 10^{-4}$, respectively; and with $c_1 = -2, c_2 = -4$, the errors are $4.1 \times 10^{-3}$, $1.0 \times 10^{-3}$, and $6.6 \times 10^{-4}$, respectively. Again the accuracy of the continuum model increases with $N_1$, $N_2$, and $M$.

## V. Conclusion and Future Work

In this paper we analyze the convergence of a sequence of Markov chains to its continuum limit, the solution of a PDE, in a two-step procedure. We provide precise sufficient conditions for the convergence and the explicit rate of the convergence. Based on such convergence we approximate the Markov chain modeling a large wireless sensor network by a nonlinear diffusion-convection PDE.

With the sophisticated mathematical tools available for PDEs, this approach provides a framework to model and simulate networks with a very large number of components, which is practically infeasible for Monte Carlo simulation. Such a

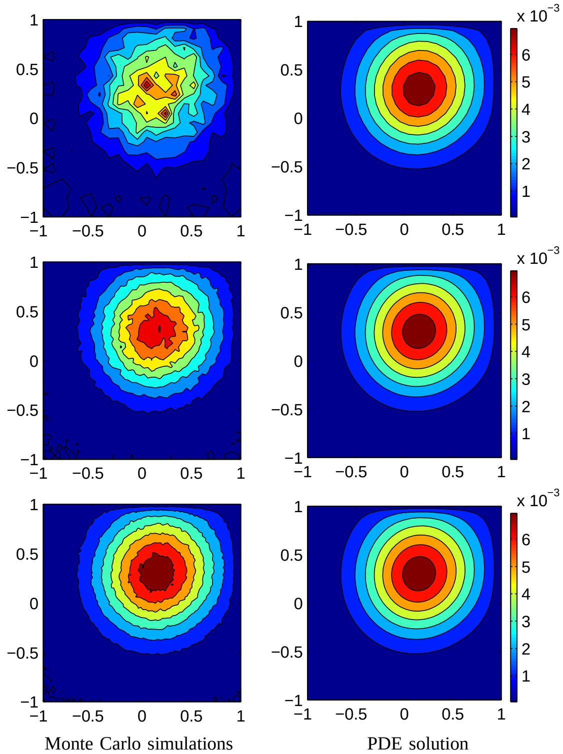


Fig. 10. The Monte Carlo simulations (from top to bottom, with $N_1 = N_2 = 20, 50, 80$, respectively, and $M = N_1^3$) and the PDE solution of a two-dimensional network, with $b_1 = b_2 = 1/4$ and $c_1 = -2, c_2 = -4$, at $t = 0.1s$.

tool enables us to tackle problems such as performance analysis and prototyping, resource provisioning, network design, network parametric optimization, network control, network tomography, and inverse problems, for very large networks. For example, we can now use the PDE model to optimize some performance metric of a large network by adjusting the placement of destination nodes or the routing parameters (coefficients in convection terms), with relatively negligible computational overhead compared with that of the same task done by Monte Carlo simulation.

The approximation approach can be extended in future work with more specific considerations regarding the network, which can significantly affect the derivation of the continuum model. For example, we can seek to establish continuum models for other domains such as the Internet, cellular networks, and traffic networks; we can consider more boundary conditions other than sinks, including walls, semi-permeating walls, and their composition; the nodes could be nonuniformly located, even mobile; transmission could happen between nodes that are not immediate neighbors; and the interference between nodes could behave differently in the presence of power control.

## References


[1] R. M. Fujimoto, K. S. Perumalla, and G. F. Riley, *Network Simulation*. Morgan & Claypool Publishers, 2007.

[2] R. Bagrodia, R. Meyer, M. Takai, Y. A. Chen, X. Zeng, J. Martin, and H. Y. Song, “Parsec: a parallel simulation environment for complex systems,” *Computer*, vol. 31, no. 10, pp. 77 –85, Oct. 1998.

[3] H. Plesser, J. Eppler, A. Morrison, M. Diesmann, and M. O. Gewaltig, “Efficient parallel simulation of large-scale neuronal networks on clusters of multiprocessor computers,” in *Euro-Par 2007 Parallel Processing*.

[4] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Processes, with Applications to Stochastic Systems Theory*. Cambridge, MA: MIT Press, 1984.

[5] R. Norberg, “Anomalous PDEs in Markov chains: Domains of validity and numerical solutions,” *Finance and Stochastics*, vol. 9, no. 4, pp. 519–537, October 2005. [Online]. Available: http://ideas.repec.org/a/spr/finsto/v9y2005i4p519-537.html

[6] R. W. R. Darling and J. R. Norris, “Differential equation approximations for Markov chains,” *Probability Surveys*, vol. 5, p. 37, 2008. [Online]. Available: doi:10.1214/07-PS121

[7] E. K. P. Chong, D. Estep, and J. Hannig, “Continuum modeling of large networks,” *Int. J. Numer. Model.*, vol. 21, no. 3, pp. 169–186, 2008.

[8] S. L. Sobolev, *Partial Differential Equations of Mathematical Physics*. Courier Dover Publications, 1964.

[9] R. G. Mortimer, *Mathematics for Physical Chemistry*. Academic Press, 2005.

[10] M. Gillman, *An Introduction to Mathematical Models in Ecology and Evolution: Time and Space*. Wiley-Blackwell, 2009.

[11] T. Hens and M. O. Rieger, *Financial Economics*. Springer, 2010.

[12] G. R. Liu and S. S. Quek, *The Finite Element Method: A Practical Course*. Butterworth-Heinemann, 2003.

[13] A. R. Mitchell and D. F. Griffiths, *The Finite Difference Method in Partial Differential Equations*. Wiley, 1980.

[14] E. W. C. v. Groesen and J. Molenaar, *Continuum Modeling in the Physical Sciences*. Society for Industrial and Applied Mathematics, 2007.

[15] H. B. Mühlhaus, *Continuum Models for Materials with Microstructure*. Wiley, 1995.

[16] W. F. Phillips, *A New Continuum Model for Traffic Flow*. U.S. Dept. of Transportation, Research and Special Programs Administration National Technical Information Service [distributor], 1981.

[17] D. Grünbaum, “Translating stochastic density-dependent individual behavior with sensory constraints to an Eulerian model of animal swarming,” *J Math Biol*, vol. 33, pp. 139–161, 1994.

[18] J. M. Harrison, “Heavy traffic analysis of a system with parallel servers: Asymptotic optimality of discrete-review policies,” *The Annals of Applied Probability*, vol. 8, no. 3, pp. pp. 822–848, 1998. [Online]. Available: http://www.jstor.org/stable/2667208

[19] J. G. Dai and J. M. Harrison, “Reflecting brownian motion in three dimensions: A new proof of sufficient conditions for positive recurrence,” *Mathematical Methods of Operations Research*, 2009.

[20] M. Bramson, J. G. Dai, and J. M. Harrison, “Positive recurrence of reflecting Brownian motion in three dimensions,” *ArXiv e-prints*, Sep. 2010.

[21] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *Information Theory, IEEE Transactions on*, vol. 46, no. 2, pp. 388–404, Mar 2000.

[22] E. W. Grundke and A. N. Z. Heywood, “A uniform continuum model for scaling of ad hoc networks,” in *ADHOC-NOW*, 2003, pp. 96–103.

[23] R. Bakhshi, L. Cloth, W. Fokkink, and B. R. Haverkort, “Meanfield analysis for the evaluation of gossip protocols,” *SIGMETRICS Perform. Eval. Rev.*, vol. 36, pp. 31–39, November 2008. [Online]. Available: http://doi.acm.org/10.1145/1481506.1481513

[24] M. E. J. Newman, C. Moore, and D. J. Watts, “Mean-field solution of the small-world network model,” *Physical Review Letters*, vol. 84, pp. 3201–3204, Apr. 2000.

[25] R. B. Guenther and J. W. Lee, *Partial Differential Equations of Mathematical Physics and Integral Equations*. Mineola, NY: Courier Dover Publications, 1996.

[26] T. C. Gard, *Introduction to Stochastic Differential Equations (Pure and Applied Mathematics)*. Marcel Dekker Inc, 1987.

[27] H. J. Kushner and G. G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[28] P. Billingsley, *Probability and Measure*. New York, NY: Wiley, 1995.